On the convergence of Optimized Schwarz Methods by way of Matrix Analysis

Sébastien Loisel[1] and Daniel B. Szyld[1]

Temple University {szyld|loisel}@temple.edu

Summary. Domain decomposition methods are widely used to solve in parallel large linear systems of equations arising in the discretization of partial differential equations. Optimized Schwarz Methods (OSM) have been the subject of intense research because they lead to algorithms that converge very quickly. The analysis of OSM has been a very challenging research area and there are currently no general proofs of convergence for the optimized choices of the Robin parameter in the case of overlap. In this article, we apply a proof technique developed for the analysis of Schwarz-type algorithms using matrix analysis techniques and specifically using properties of matrix splittings, to the Optimized Schwarz algorithms. We thus obtain new general convergence results, but they apply only to large Robin parameters, which may not be the optimal ones.

1 Introduction

Schwarz iterative methods for the solution of boundary value problems have been extensively studied; see, e.g., [12], [14], [16]. When some of these methods are used as parallel preconditioners, they have been shown to scale perfectly in many thousands of processors; see, e.g., [6].

The main idea is to split the overall domain Ω into multiple subdomains $\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_p$, then solve Dirichlet problems on each subdomain, and iterate. This is usually refered to as multiplicative Schwarz methods. There are many variants, which we name after the type of boundary condition used on the artificial interfaces, for instance Schwarz-Neumann or Schwarz-Robin. While the classical Schwarz methods (with Dirichlet boundary conditions on the artificial interface) are very well understood, [12], [14], [16], less progress has been made in the theory of Schwarz-Robin methods. In this paper, we present a new convergence result for large Robin parameters, which may be useful in certain situations.

One difficulty with Schwarz-Robin algorithms is the choice of the real parameter α for the differential operator $\alpha u + D_\nu u$ (D_ν is a normal derivative) on the artificial interface. The first analysis of a Schwarz-Robin method was performed with some generality by Lions [9] for the case of zero overlap. It is well-known that overlap usually improves the convergence rate of Schwarz algorithms. Detailed studies for the overlap case do exist for the case of simple domains, such as rectangles, half planes, or hemispheres, and the main analytical tool is the use of Fourier transforms. For a history, review, analysis and extensive bibliography of such methods, see [5]. For general domains, and two overlapping subdomains, Kimn [7], [8], proved convergence of the method under certain conditions; see also [10].

The main contribution of this paper is to show convergence of the (multiplicative) Schwarz-Robin iteration for elliptic problems on p general subdomains of a general two-dimensional domain, when α is sufficiently large.

For classical Schwarz methods, algebraic representations have been proposed; see [4], [1], [11], and references therein. Our approach is inspired by some of these papers, and by [7], [8], and [15], and it was prompted by some general matrix analytic convergence results in the very recent paper [3] based on the theory of splittings. We are able to apply these new results using an inverse trace inequality, which we prove in Lemma 1 in section 2.

1.1 Model problem and notation

Let Ω be an open region in the plane. For simplicity, we consider the problem

$$\Delta u = f \text{ in } \Omega, \text{ with } u = 0 \text{ on } \partial\Omega; \tag{1}$$

although our proof holds, mutatis mutandis, for any symmetric and coercive elliptic operator.

We use a piecewise linear finite element discretization of (1). We denote by $T = T_h$ the triangulation of the finite element method (which depends on the mesh parameter h), and by v_i its vertices.

For the description of the discretized problem, we abuse the notation, and call u the vector of nodal values which approximate the function u at the nodes v_i. The discretized problem is then

$$Au = f,$$

where the $n \times n$ matrix A has entries

$$A_{ij} = \int_\Omega \nabla\phi_i \cdot \nabla\phi_j,$$

with ϕ_i and ϕ_j being piecewise linear basis functions corresponding to vertices v_i and v_j of the finite element discretization. The finite element space is denoted $V_{h,0}(\Omega) \subset H_0^1(\Omega)$.

We now define Schwarz and Optimized Schwarz algorithms. First, we introduce the notion of *restriction matrices.* This is a matrix R of the form

$$R = \begin{bmatrix} 0\,0\,1\,0\,0\,0 \\ 0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,0 \end{bmatrix},$$

i.e, the $m \times n$ matrix R is formed by taking m rows from an $n \times n$ permutation matrix. The transpose R^T is known as a prolongation matrix. Note that $RR^T = I_m$, the identity in the m-dimensional space.

Given a *domain decomposition* $\Omega = \Omega_1 \cup \ldots \cup \Omega_p$, such that each Ω_i is a union of triangles in T_h, we can form restriction matrices $R_1, \ldots, R_p$ which restrict to those vertices in the *interior* of Ω_i. These matrices are uniquely determined up to permutation of the rows.

Let $u_0 \in V_0(\Omega)$. The (multiplicative) Schwarz iteration can be phrased algebraically as

$$u_{k+\frac{j}{p}} = u_{k+\frac{j-1}{p}} + M_j(f - Au_{k+\frac{j-1}{p}}) \text{ for } k = 0, 1, 2, \ldots \text{ and } j = 1, 2, \ldots, p.$$

This recurrence relation can be concisely stated in terms of the error terms $u_k - u$:

$$u_{k+1} - u = (I - M_pA) \ldots (I - M_1A)(u_k - u) \text{ for } k = 0, 1, 2, \ldots, \quad (2)$$

where $M_i = R_i^T(R_iAR_i^T)^{-1}R_i$, $i = 1, \ldots, p$. If we define $A_i = R_iAR_i^T$, we can write instead $M_i = R_i^T A_i^{-1} R_i$. If the subdomain Ω_i has n_i vertices in its interior, then A_i is an $n_i \times n_i$ matrix with entries

$$A_{i,jk} = \int_{\Omega_i} \nabla\phi_j \cdot \nabla\phi_k,$$

where $\phi_1, \ldots, \phi_{n_i}$ are the piecewise linear basis functions of $V_0(\Omega_i)$. Note that A_i is the finite element discretization of a Dirichlet problem in Ω_i. The main idea of Optimized Schwarz Methods is to use Robin problems on the subdomains, instead of Dirichlet problems. This means that we must replace A_i by a matrix $\widetilde{A}_i$ of a Robin problem. The FEM discretization of a Dirichlet problem, (e.g., A_i) does not have any degrees of freedom along the boundary Ω_i (which is why our restriction matrices correspond to the vertices in the interior of Ω_i). However, the FEM discretization of a Robin problem contains degrees of freedom along the boundary. We want to keep the same matrices R_i for both algorithms, which means that the domain decomposition $\Omega_1^\circ, \ldots, \Omega_p^\circ$ for the Robin version is not the same as the domain decomposition $\Omega_1, \ldots, \Omega_p$ of the Dirichlet version.

For $i = 1, \ldots, p$, define Ω_i° to be the set of triangles in Ω_i which do not have a vertex on $\overline{\partial\Omega_i \setminus \partial\Omega}$. The matrix $\widetilde{A}_i$ of the Neumann problem for Ω_i° is

$$\widetilde{A}_{i,jk} = \int_{\Omega_i^\circ} \nabla\phi_j \cdot \nabla\phi_k,$$

where $\phi_1, \ldots, \phi_{n_i}$ are the piecewise linear basis functions of $V_0(\Omega_i)$. The matrix of the Robin problem with real parameter $\alpha > 0$ is then $\widetilde{A}_i + \alpha B_i$, where

$$B_{i,jk} = \int_{\partial \Omega_i^\circ} \phi_j \phi_k.$$

Optimized Schwarz algorithms use the matrices

$$\widetilde{M}_i = R_i^T (\widetilde{A}_i + \alpha B_i)^{-1} R_i \tag{3}$$

instead of $M_i = R_i^T A_i^{-1} R_i$, and the iteration is then

$$u_{k+1} - u = (I - \widetilde{M}_p A) \dots (I - \widetilde{M}_1 A)(u_k - u) \text{ for } k = 0, 1, 2, \dots \tag{4}$$

A related algorithm uses a coarse grid correction. This is achieved by choosing R_0 to be an averaging operator of the form

$$R_0 = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.25 & 0.4 & 1 & 0.15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.28 & 1 \end{bmatrix};$$

i.e., an $n_0 \times n$ matrix with non-negative entries. Usually, we choose R_0 so that R_0^T is the matrix of an interpolation operator from a coarse grid T_H to the fine grid T_h, where $H \gg h$. We then let $M_0 = R_0^T A_0^{-1} R_0$, where $A_0 = R_0 A R_0^T$ and we use $\widetilde{M}_i$ for $i = 1, \dots, p$. The coarse grid corrected OSM then fits our framework and the iteration is

$$u_{k+1} - u = (I - \widetilde{M}_p A) \dots (I - \widetilde{M}_1 A)(I - M_0 A)(u_k - u) \text{ for } k = 0, 1, 2, \dots \tag{5}$$

As we refine the mesh, the triangulation T_h changes. We may also want to change the domain decomposition $\Omega_1, \dots, \Omega_{p(h)}$, increasing the number of subdomains $p(h)$ as h goes to zero, in such a way that the amount of work per subdomain remains constant.

2 Convergence of OSM

In this section, we prove that the OSM (4) converges. We further assume that the triangulation T_h is quasi-uniform; see, e.g., [2].

Lemma 1 (Inverse Trace Inequality). *Let Ω be a domain with a quasi-uniform triangulation $T = T_h$. Let U be a set of triangles in T and let Γ be a set of edges in T. Let $V(\Omega, \Gamma, U)$ denote the space of piecewise linear functions on T which are zero at every vertex of U outside of Γ. Then, there is a constant $C < \infty$ which depends on the regularity parameters of T_h, but not on h, Γ, U or u, such that*

$$\int_U (\nabla u)^2 \leq C h^{-1} \int_\Gamma u^2, \tag{6}$$

for every $u \in V(\Omega, \Gamma, U)$.

Proof. What makes this inverse trace inequality possible is that the quadratic form $\int_U (\nabla u)^2$ depends only on the function values of u at the vertices of Γ. Indeed, let A_U be the matrix of $\int_U (\nabla u)^2$ on the space $V(\Gamma, U)$, with entries $A_{U,ij} = \int_U \nabla\phi_i \cdot \nabla\phi_j$ when both v_i and v_j are on Γ, and with $A_{U,ij} = 0$ otherwise. We introduce the "vertexwise" norm $u^T u = \sum_{v_i \in \Gamma} u^2(v_i)$. The matrix A_U has the same sparsity pattern, or is sparser than the matrix A of the elliptic problem. Hence, the bandwidth of A_U is independent of Γ and U, but does depend on the regularity parameters of T_h. By the regularity of T_h as $h \to 0$, A (and hence A_U) has a bandwidth $N(h) < N < \infty$, uniformly in h. By [13, §6.3.2], the entries of A_U are bounded by some constant C_1 which is independent of h. Hence, $u^T H u \le C_2 u^T u$, where C_2 depends only on the regularity parameters of T_h. Likewise, by a variant of [2, §6.2], there is a C_3 such that $u^T u \le h^{-1} C_3 \int_\Gamma u^2$. Putting these together, we obtain (6).

To prove our main result, we must first cite a Theorem from [3], where the notation M^H stands for the conjugate transpose of M.

Theorem 1. *[3, Theorem 3.15] Let $A \in \mathbb{C}^{n\times n}$ be Hermitian and positive semidefinite. Let $\widetilde{M}_i \in \mathbb{C}^{n\times n}$, $i = 1, \dots, p$, be such that*

(i) $\ker A\widetilde{M}_i A = \ker \widetilde{M}_i A$.
(ii) There exists a number $\gamma > 1/2$ such that

$$\widetilde{M}_i + \widetilde{M}_i^H - 2\gamma \widetilde{M}_i^H A \widetilde{M}_i \tag{7}$$

is positive semidefinite on the range of A, for $i = 1, \dots, p$.
(iii) $\cap_{i=1}^p \ker \widetilde{M}_i A = \ker A$.

Then the iteration (4) converges for any initial vector u_0.

We note that in our case, the matrices are real and symmetric and therefore $\widetilde{M}_i^H = \widetilde{M}_i$. We also note that Theorem 1 is general and does not require that the matrices $\widetilde{M}_i$ have the particular form discussed in this paper.

Theorem 2. *There is a constant C, which depends only on the regularity parameters of T_h, such that for all $\alpha \ge Ch^{-1}$, the OSM iteration defined by (4) converges for any initial vector u_0.*

Proof. We verify the hypotheses of Theorem 1. Since A is injective, part (i) is automatically satisfied. Part (iii) is also easily checked: it suffices to show that $\cap_{i=1}^p \ker \widetilde{M}_i = \{0\}$. Let f be such that $\widetilde{M}_i f = 0$ for $i = 1, \dots, p$. Then, $R_i f = (\widetilde{A}_i + \alpha B_i)0 = 0$ for $i = 1, \dots, p$. Hence, $f|_{\Omega_i} = 0$ for $i = 1, \dots, p$. Since $\Omega = \Omega_1 \cup \ldots \cup \Omega_p$, we have that $f = 0$.

For part (ii), we multiply (7) on the left by $(\widetilde{A}_i + \alpha B_i) R_i$ and on the right by its transpose $R_i^T(\widetilde{A}_i + \alpha B_i)$, using the expression of the symmetric matrix $\widetilde{M}_i$, and that $R_i R_i^T = I_{n_i}$, then it follows that all we need to show is that

$u^T(\widetilde{A}_i + \alpha B_i - \gamma A_i)u \geq 0$ for every $u \in V_0(\Omega_i)$. We show this now for $\gamma = 1$. We calculate

$$u^T(\widetilde{A}_i + \alpha B_i - A_i)u = \int_{\Omega_i^\circ} (\nabla u)^2 + \alpha \int_{\partial\Omega_i^\circ} u^2 - \int_{\Omega_i} (\nabla u)^2$$
$$= \alpha \int_{\partial\Omega_i^\circ} u^2 - \int_{U_i} (\nabla u)^2,$$

where $U_i = \Omega_i \setminus \Omega_i^\circ$. Since the function u is in $V_0(\Omega_i) \subset V(\Omega_i, \partial\Omega_i^\circ, U_i)$, the result follows from Lemma 1. □

Remark 1. A choice of parameter $\alpha = O(h^{-1})$ is not an optimal value (see, e.g., [5]) and will typically yield a convergence rate which is not much better than the classical Schwarz algorithm (2). In fact, in [15] a similar two-parameter variant of OSM is given, and for specific choices of these parameters which closely correspond to our choice of $\alpha = O(h^{-1})$, the method coincides with classical Schwarz; see equation (4.3) of [15].

3 Numerical experiments

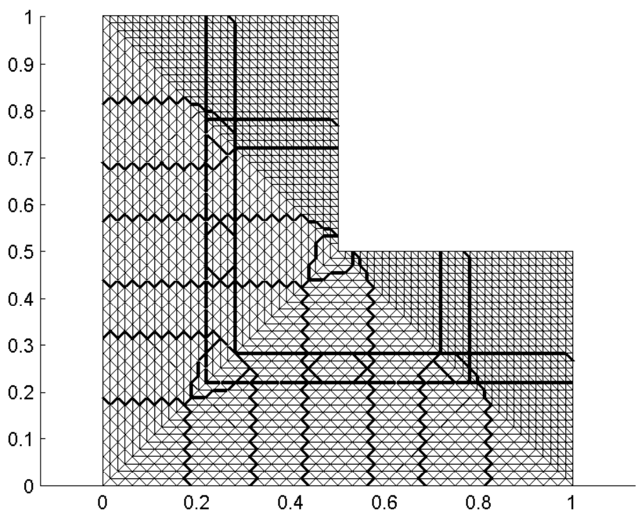

Fig. 1. L-shaped domain subdivided into twelve overlapping subdomains

We present a numerical experiment for the problem (1) on an L-shaped region split into twelve subdomains, and we have $h = 1/32$. We use piecewise linear elements. The domain, the subdomains and the mesh are depicted in

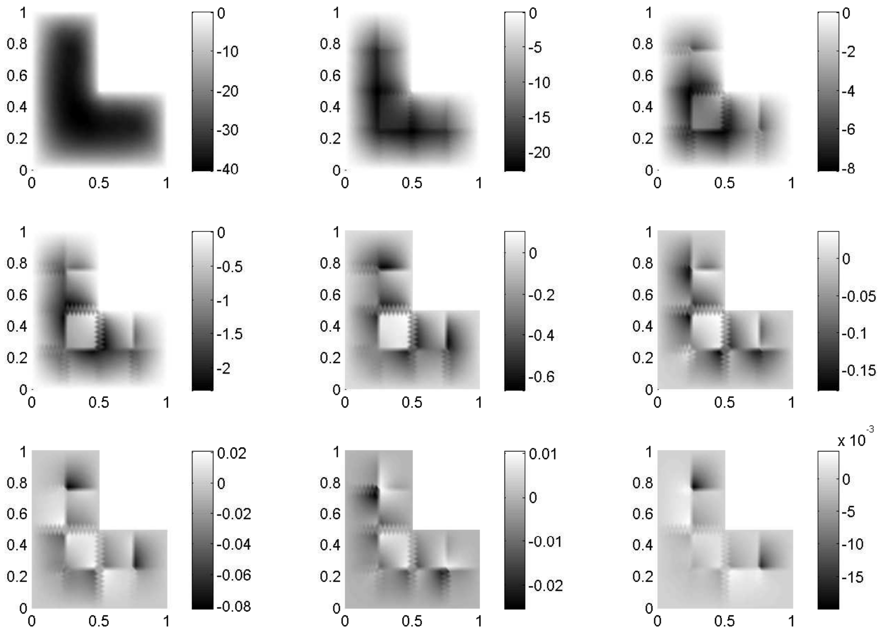

Fig. 2. Error terms of the same nine iterates with twelve subdomains.

Figure 1. In Figure 2 we show the error terms of the first 9 iterations of the OSM algorithm using $\alpha = 5$ for these 12 subdomains. Note the scale in each of the error iterates.

References

[1] M. Benzi, A. Frommer, R. Nabben, and D. B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605-639, 2001.

[2] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, Second Edition. Springer-Verlag, New York, 2002.

[3] A. Frommer, R. Nabben, and D. B. Szyld. Convergence of stationary iterative methods for Hermitian semidefinite linear systems and applications to Schwarz methods. *SIAM Journal on Matrix Analysis and Applications*, 30:925–938 (2008).

[4] A. Frommer and D. B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

[5] M. J. Gander. Optimized Schwarz Methods. *SIAM Journal on Numerical Analysis*, 44:699-731, 2006.
[6] D. Keyes. Domain Decomposition Methods in the Mainstream of Computational Science. In *Domain Decomposition Methods in Science and Engineering.* Fourteenth International Conference on Domain Decomposition Methods, Cocoyoc, Mexico, I. Herrera, D. E. Keyes, O. B. Widlund, and R. Yates, editors. Press of the Universidad Nacional Autónoma de México (UNAM), Mexico, 2003, pp. 79–94.
[7] J.-H. Kimn. Overlapping Schwarz Algorithms using Discontinuous Iterates for Poisson's Equation. Ph.D. Thesis, Department of Mathematics, New York University, New York, 2001.
[8] J.-H. Kimn. A convergence theory for an overlapping Schwarz algorithm using discontinuous iterates. *Numerische Mathematic*, 100:117-139, 2005.
[9] P. L. Lions. On the Schwarz alternating method. III: A variant for nonoverlapping subdomains. In T. F. Chan, R. Glowinski, J. Périaux, O. Widlund, eds. *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, held in Houston, Texas, March 20-22, 1989. SIAM, 1990, Philadelphia, pp. 202–223.
[10] S. Loisel and D. B. Szyld. On the convergence of Algebraic Optimizable Schwarz Methods with applications to elliptic problems. Research Report 07-11-16, Department of Mathematics, Temple University, November 2007.
[11] R. Nabben and D. B. Szyld. Schwarz iterations for symmetric positive semidefinite problems. *SIAM Journal on Matrix Analysis and Applications*, 29:98–116, 2006.
[12] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations.* Oxford Science Publication, Clarendon Press, Oxford, 1999.
[13] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations.* Springer, Berlin, 1994. Second, Corrected Printing 1997.
[14] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, Cambridge, New York, Melbourne, 1996.
[15] A. St-Cyr, M. J. Gander and S. J. Thomas. Optimized multiplicative, additive and restricted additive Schwarz preconditioning. *SIAM Journal on Scientific Computing*, 29:2402-2425, 2007.
[16] A. Toselli and O. Widlund. *Domain Decomposition Methods – Algorithms and Theory.* Springer Series in Computational Mathematics 34, Springer, Berlin, Heidelberg, New York, 2005.